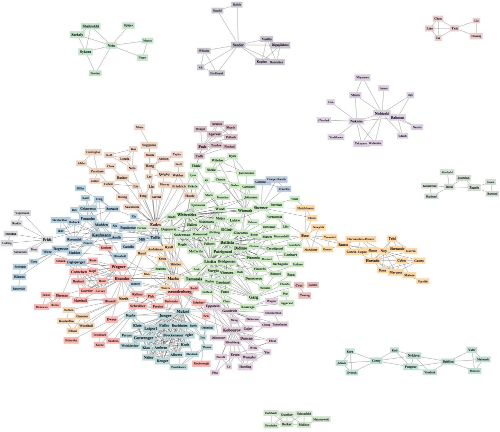
Maximum Differential Graph Coloring

Sankar Veeramoni
University of Arizona
Joint work with Yifan Hu at AT&T Research and Stephen Kobourov at University of Arizona

Outline

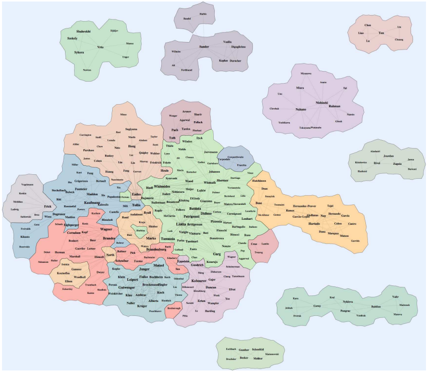
Motivation

- Traditionally, relational data are visualized as graphs.
- Points and labels are colored based on the clustering.



Motivation

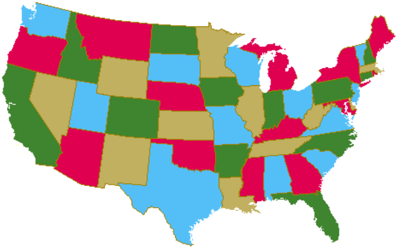
- Gmap : Visualize high dimensional data as maps
- Clear clustering information
- Explicit boundary for the clusters.
- Assign colors to clusters of map
  - visually appealing and informative



Outline

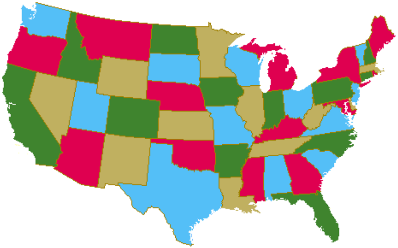
Map Coloring Problem

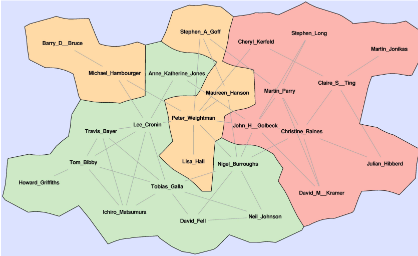
- Four Color Theorem



Map Coloring Problem

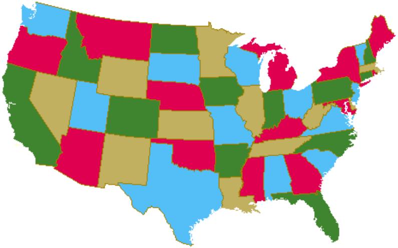
- Four Color Theorem
    - Countries should be contiguous

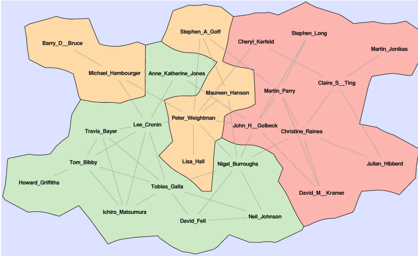




Map Coloring Problem

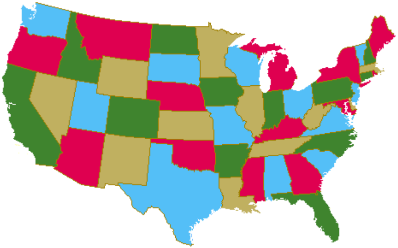
- Four Color Theorem
  - Countries should be contiguous
- Our "countries" may not be contiguous
  - Cluster Fragmentaion
  - label overlap removal

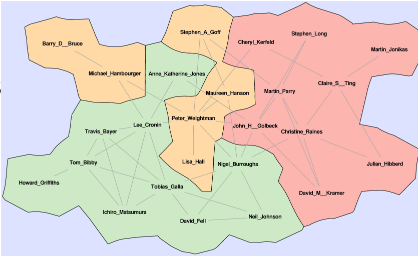




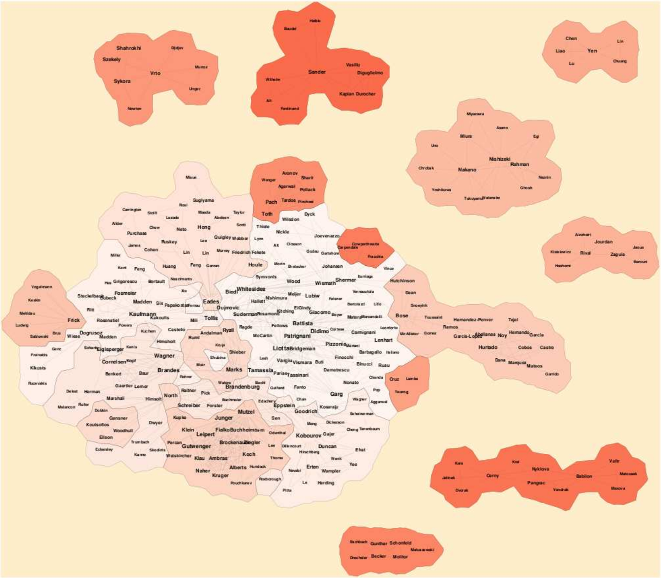
Map Coloring Problem

- Four Color Theorem
  - Countries should be contiguous
- Our "countries" may not be contiguous
  - Cluster Fragmentaion
  - label overlap removal
- # colors = #countries
- Careful with colors for adjacent countries

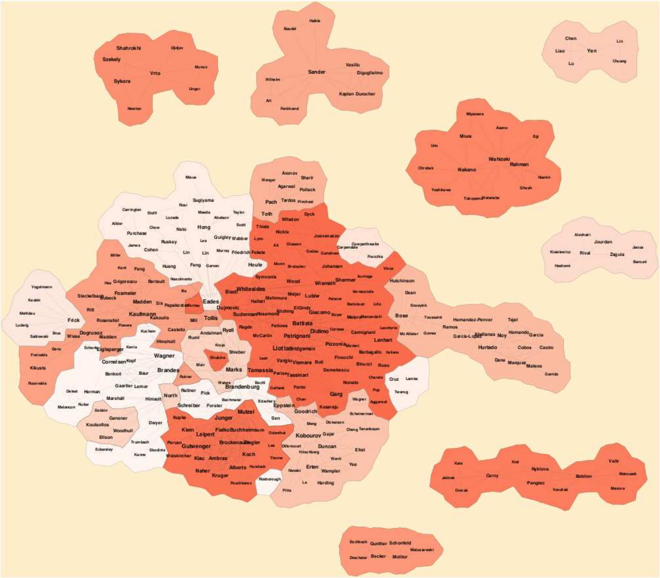




Bad Coloring

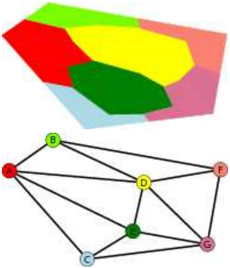


Good Coloring



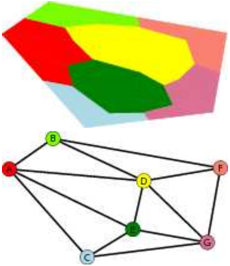
Problem Overview

- Turn into a graph coloring problem
- Create country graph $G_c = (V_c, E_c)$



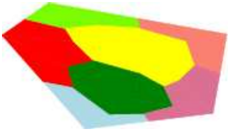
Problem Overview

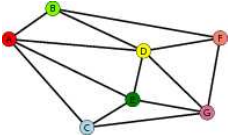
- Turn into a graph coloring problem
- Create country graph $G_c = (V_c, E_c)$
- vertices in V_c are countries
- $(i, j) \in E_c \iff$ countries i and j are neighbours



Problem Overview

- Turn into a graph coloring problem
- Create country graph $G_c = (V_c, E_c)$
- vertices in V_c are countries
- $(i, j) \in E_c \iff$ countries i and j are neighbours
- Assign colors to nodes of G

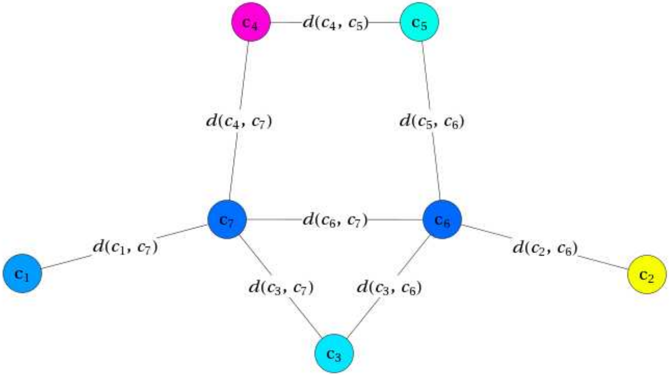




The Graph Coloring Problem

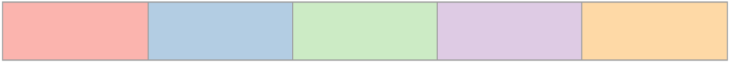
- How to color vertices to maximize differences along the edges?

$$\max_{c \in C} \min_{\{i,j\} \in E_c} w_{i,j} d\left(c_i, c_j\right), \quad C\text{: } \textit{the color space}$$

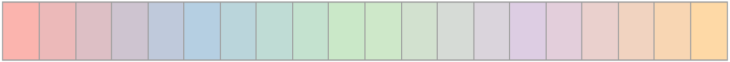


Color distance

- Color selection
  - “map-like” pastel colors
  - 5 color palette from ColorBrewer



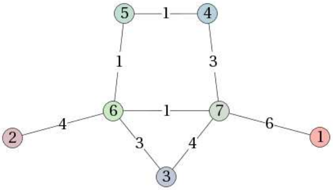
  - Blend them to get as many colors as needed

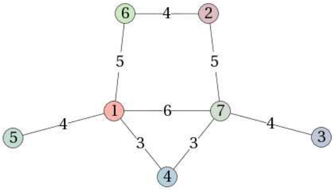


  - Discrete color space; any two consecutive colors are very similar
- Final result is to be printed in black and white.
- The color space is strictly 1D.

Maximum differential graph coloring problem

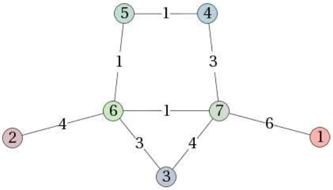
- Colors form a line in the color space
- Label nodes with numbers from $\{1, 2, \ldots, |V|\}$
- Maximize the label difference along the edges in the graph.

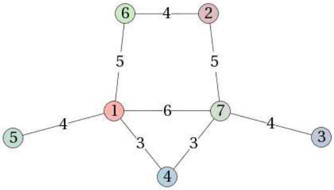




Maximum differential graph coloring problem

- Colors form a line in the color space
- Label nodes with numbers from $\{1, 2, \ldots, |V|\}$
- Maximize the label difference along the edges in the graph.





- Maximum differential graph coloring is a Bijection $c : V \rightarrow \{1, 2, \ldots, |V|\}$ that optimize

$$\max_{c \in C} \min_{\{i,j\} \in E} w_{ij}|c(i) - c(j)| \quad (1)$$

$w_{ij} \rightarrow$ edge weight.we assume $w_{ij} = 1$.
$C = \{all\ permutations\ of\ \{1, 2, \cdots |V|\}\}$.

Outline

Related Work

- The complementary problem : Minimize the labeling differences along the edges.

Related Work

- The complementary problem : Minimize the labeling differences along the edges.
  - Well-studied problem in the context of minimum bandwidth

Related Work

- The complementary problem : Minimize the labeling differences along the edges.
  - Well-studied problem in the context of minimum bandwidth
  - Minimum bandwidth is NP-complete, with a reduction from 3SAT
    Papadimitriou et al Computing 1975

Related Work

- The complementary problem : Minimize the labeling differences along the edges.
  - Well-studied problem in the context of minimum bandwidth
  - Minimum bandwidth is NP-complete, with a reduction from 3SAT
    Papadimitriou et al Computing 1975
  - NP-complete to find any constant approximation
  - Effective heuristics have been proposed

Related Work

- The complementary problem : Minimize the labeling differences along the edges.
  - Well-studied problem in the context of minimum bandwidth
  - Minimum bandwidth is NP-complete, with a reduction from 3SAT
    Papadimitriou et al Computing 1975
  - NP-complete to find any constant approximation
  - Effective heuristics have been proposed
- Anti Bandwidth Problem : maximize the labeling differences along the edges.

Related Work

- The complementary problem : Minimize the labeling differences along the edges.
  - Well-studied problem in the context of minimum bandwidth
  - Minimum bandwidth is NP-complete, with a reduction from 3SAT
    Papadimitriou et al Computing 1975
  - NP-complete to find any constant approximation
  - Effective heuristics have been proposed
- Anti Bandwidth Problem : maximize the labeling differences along the edges.
  - Problem same as Maximum differential graph coloring problem

Related Work

- The complementary problem : Minimize the labeling differences along the edges.
  - Well-studied problem in the context of minimum bandwidth
  - Minimum bandwidth is NP-complete, with a reduction from 3SAT
    Papadimitriou et al Computing 1975
  - NP-complete to find any constant approximation
  - Effective heuristics have been proposed
- Anti Bandwidth Problem : maximize the labeling differences along the edges.
  - Problem same as Maximum differential graph coloring problem
  - Probelem is NP-Complete, with a reduction from Hamiltonian Path
    Leung et al Computing 1984

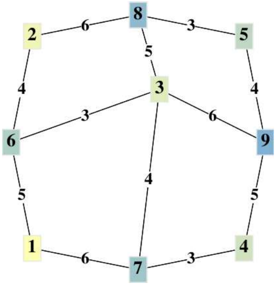
Outline

k-differential colorable

- A *k-differential coloring* of G is one in which the absolute coloring difference of the endpoints for any edge is k or more.

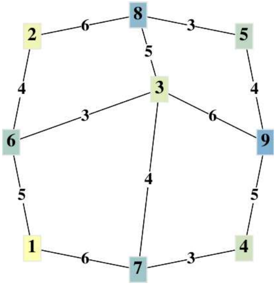
k-differential colorable

- A *k-differential coloring* of G is one in which the absolute coloring difference of the endpoints for any edge is k or more.
- 3 differential coloring of a 3 $\times$ 3 grid.



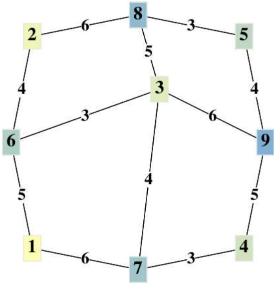
k-differential colorable

- A *k-differential coloring* of G is one in which the absolute coloring difference of the endpoints for any edge is k or more.
- 3 differential coloring of a 3 $\times$ 3 grid.



k-differential colorable

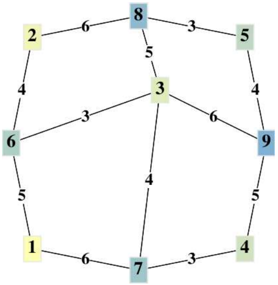
- A *k-differential coloring* of G is one in which the absolute coloring difference of the endpoints for any edge is k or more.
- 3 differential coloring of a 3 $\times$ 3 grid.



- differential chromatic number : $dc(G) = k \Rightarrow$ G is k-differential colorable, but not (k + 1)-differential colorable.

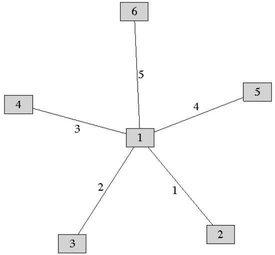
k-differential colorable

- A *k-differential coloring* of G is one in which the absolute coloring difference of the endpoints for any edge is k or more.
- 3 differential coloring of a 3 $\times$ 3 grid.



- differential chromatic number : $dc(G) = k \Rightarrow$ G is k-differential colorable, but not (k + 1)-differential colorable.

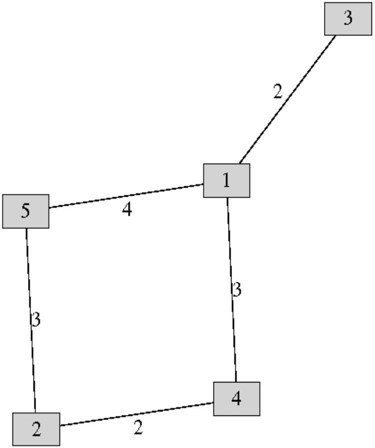
- Star Graph always have differential chromatic number $= 1$

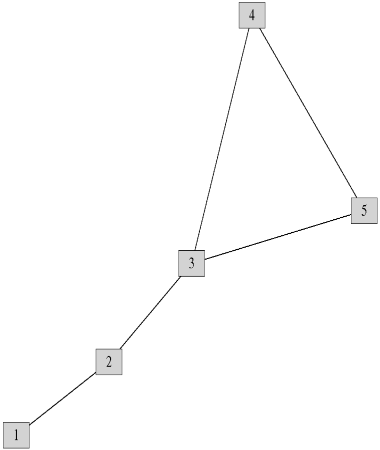


Complement of a Graph

- The complement of graph G,

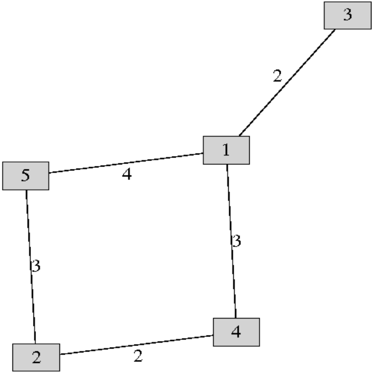
$$\bar{G} = \{V, \bar{E}\}, \text{ where } \bar{E} = \{\{i,j\} | i \neq j, i,j \in V, \text{ and } \{i,j\} \notin E\}.$$

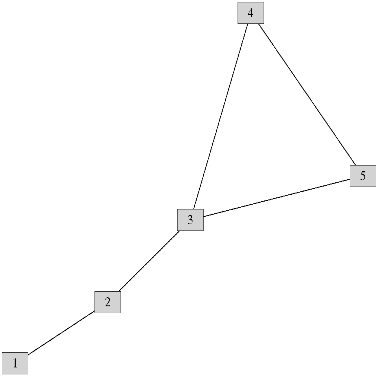




2 differential coloring is NP-Complete

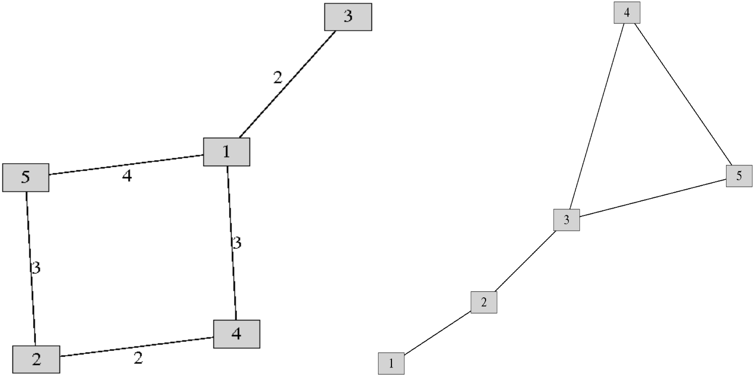
- A 2 differential coloring of a graph G exists iff $\bar{G}$ is Hamiltonian graph.





2 differential coloring is NP-Complete

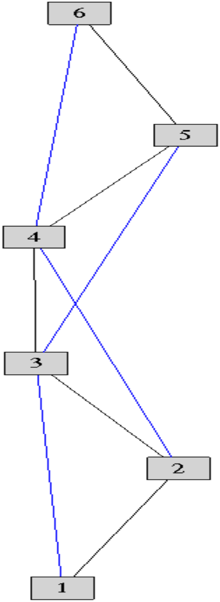
- A 2 differential coloring of a graph G exists iff $\bar{G}$ is Hamiltonian graph.



- Finding whether a graph has a Hamiltonian path is NP-complete.
- Therefore finding whether a graph is 2-differential colorable is NP-complete!

k hamiltonian path

- A *k-Hamiltonian path* ($k \geqslant 1$) of G is a Hamiltonian path, such that each i-th node on the path is connected to the j-th node, if $|i - j| \leqslant k$.
- 2 hamiltonian Graph
- A k-differential coloring ($k \geqslant 2$) of a graph exists iff a $(k-1)$-Hamiltonian path exists in the complement.



Outline

Exact Algorithm

- The equivalence between k-differential coloring and $(k-1)$-Hamiltonian path means that
- Find a k-Hamiltonian path of $\bar{G} \rightarrow G$ is $(k+1)$-colorable.
- An exact algorithm: kpath
  - start from a node
  - tries to add a neighbor to the last node in the path, and checks that this maintains the k-path condition.
  - if the condition is violated, the next neighbor is explored, or the algorithm back tracks.
- But the complexity is exponential.
- For large graphs, need heuristic algorithms.

Heuristic to find differential chromatic number

- Discrete maxmin coloring problem was converted to continuous maximization problem.

$$\max \sum_{\{i,j\} \in E} w_{ij}(c_i - c_j)^2, \text{ subject to } \sum_{i \in V} c_i^2 = 1 \quad (2)$$

where $c \in R^{|V|}$.

- Solution : eigenvector corresponding to the largest eigenvalue of the weighted Laplacian of the country graph
- Greedy refinement algorithm : repeatedly swaps pairs of vertices
- We call this algorithm GSpectral (Greedy Spectral).
- Running Time $= O(|E|^2)$.

Differential chromatic number given by GSpectral

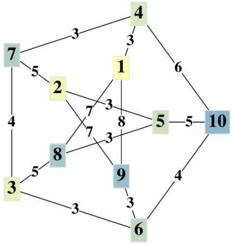


Figure: Petersen graph. dc(G) = 3. Gspectral(G) = 3

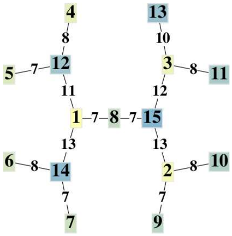


Figure: Binary tree graph. dc(G) = 7. Gspectral(G) = 5

Differential chromatic number given by GSpectral

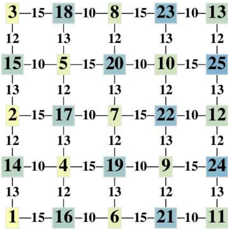


Figure: Grid 5 $\times$ 5. dc(G) = 10. Gspectral(G) = 7

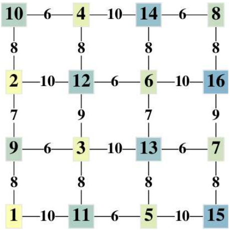


Figure: Grid 4 $\times$ 4. dc(G) = 6. Gspectral(G) = 3

Differential chromatic number given by GSpectral

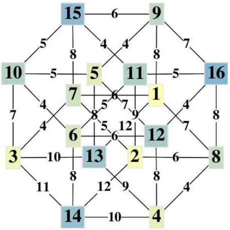


Figure: Hypercube. dc(G) = 4. Gspectral(G) = 4

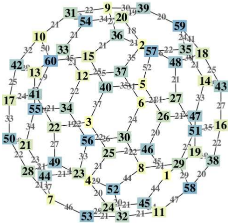


Figure: Football. dc(G) = ?. Gspectral(G) = 18

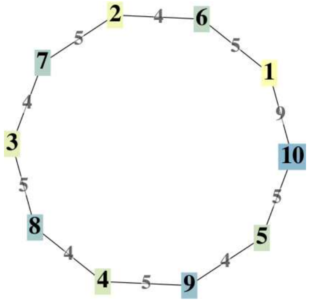
Outline

Differential Chromatic Numbers of Special Graphs

- A line graph of n nodes has differential chromatic number of $\lfloor n/2 \rfloor$.
- Consider a line graph with even number of nodes
- The labeling $n/2+1,\ 1, n/2+2,\ 2, \ldots,\ n,\ n/2$ is a $\lfloor n/2 \rfloor$ differential coloring of line graph.
- Consider the node labeled $n/2$.
- So $\lfloor n/2 \rfloor + 1$ differential coloring of line graph is not possible.

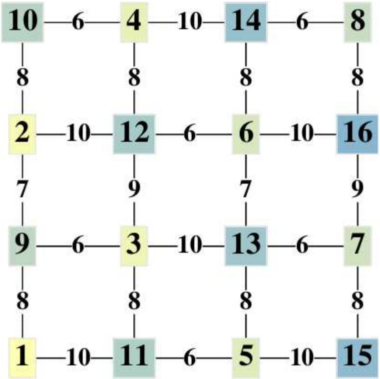
Differential Chromatic Numbers of Special Graphs

- A cycle graph of n nodes has differential chromatic number of $\lfloor (n-1)/2 \rfloor$.
- Consider a cyclic graph with even number of nodes
- The labeling $1,\ n/2+1,\ 2,\ n/2+2,\ \ldots, n/2,\ n$ is a $\lfloor (n-1)/2 \rfloor = (n/2) - 1$ differential coloring of cycle graph.
- Consider the node labeled $n/2$.
- So $n/2$ differential coloring of cycle graph is not possible.



Differential Chromatic Numbers of Special Graphs

- A grid graph of $n \times n$ nodes has differential chromatic number $= \frac{1}{2}n(n-1)$.
- If n is even, we can color the grid as show in Figure
- The formal proof of the bounds was done by Raspaud et al. Discrete mathematics'09



Outline

Future Work

- Non uniform weights based on border length.
- Improve the spectral/greedy swapping heuristic algorithm.
- Find heuristic algorithm with proven constant appriximation bound (or prove that such an algorithm dos not exist).
    - NP Hardness of approximation algorithm.